NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

GO-ESSP

Global Organization for Earth System Science Portals

NOMADS Status and GO-ESSP Vision

Glenn K. Rutledge

NOAA National Centers for Environmental Information (NCEI)

USGCRP Interagency Group on Integrated Modeling (IGIM)

GEO Climate Task Lead and GEO SBIB Board

GO-ESSP and NOMADS PI

- and -

Jordan Alpert

NOAA National Centers for Environmental Prediction

Environmental Modeling Center

NOMADS Co-I

NOMADS

The NOAA Operational Model Archive and Distribution System

Member GO-ESSP

NOMADS Data Access System

• In 2001 to overcome this deficiency in model data access, some of the Nation's top scientists actively engaged in a grass-roots framework to share data and research findings over the Internet.

• NCDC, NCEP, and GFDL initiated the NOAA Operational Model Archive and Distribution System with PMEL, NCAR, COLA, PCMDI, and others…

• Operational in NESDIS and NWS NOMADS is a distributed data service providing access to climate and weather models.

Original NOMADS Goals (c.2001)

- Establish a unified climate and weather model archive providing format independent access to retrospective models (NWP, Reanalysis, Ocean, limited Climate)
- Promote model evaluation and community feedback
- Foster research within the geo-science communities (ocean, weather, and climate) to study multiple earth systems using collections of distributed data
- Develop institutional partnerships and access via distributed open standard technologies. (GO-ESSP!)

Science Motivation

- Numerical Weather Prediction, Reanalysis and Climate model output and observations are vital to providing timely assessments of climate change and impacts:
  - climate model output statistics from NWP run in quasi-real time to identify time-dependent biases in observations;
  - assess the affect of missing data;
  - for observing network design and operation, models can be used to guide the spatial and temporal sampling frequency to resolve distributions for specific variables (temp, precip).
- Degree seeking programs heavily use NOMADS

NWS/ Real-Time NOMADS Services

- **EMC** is responsible for the enhancements, transitions-to-operations, and maintenance of more than 20 numerical prediction systems comprising NCEP's operational production suite.

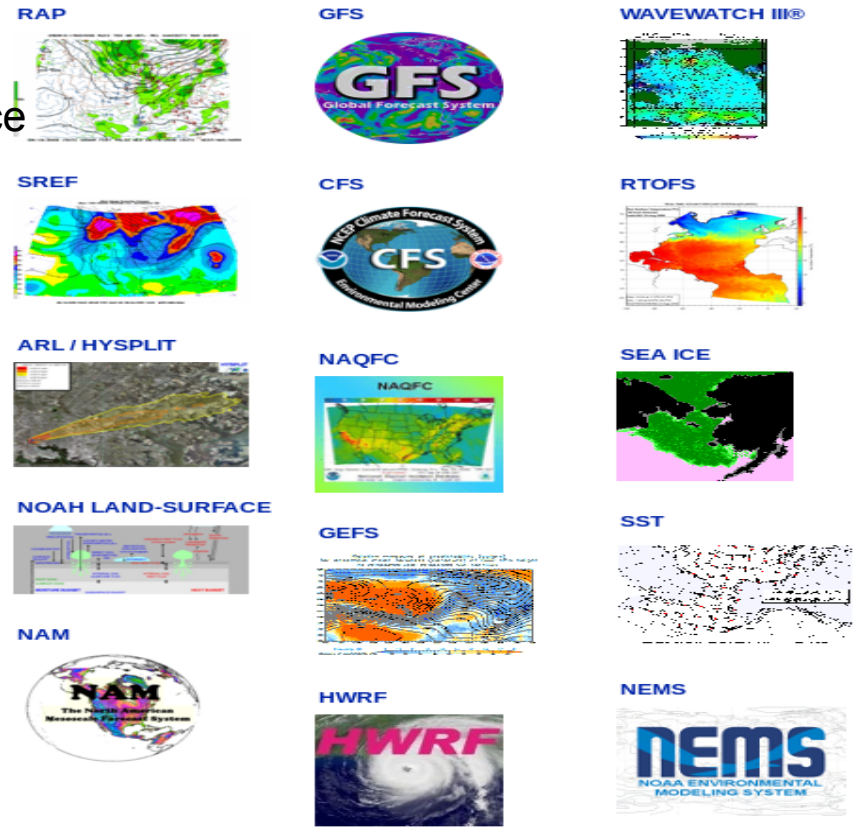
- **NOMADS real-time** services are supported by NOAA/NCEP at http://nomads.ncep.noaa.gov/.

This server provides all the real time weather service model output.

- **Access** to NCEP's operational data sets as they are being generated for up to a month in the past for some data sets (we still call this real time).

- **24x7** monitoring and data flow by NCEP staff and backup server to insure high availability for time critical access.

- **50PB** Private Archive for Model improvements



Slide: Jordan Alpert NWS/NCEP/EMC

NWS/NCEP Sub-setting Capabilities

- NCEP NOMADS services are format independent using GRADS Data Server. Applications to repackage Grib data sets:
- **"http/fast/partial"** ftp protocol allows "random access" GRIB file transfer using server side metadata description index files and a downloadable perl program that supports ftp random access. Called "http/fast/partial ftp" commands are easily scripted/automated and placed into a cron job. For examples and downloads see
  http://www.cpc.ncep.noaa.gov/products/wesley/fast_downloading_grib.html
  http://www.cpc.ncep.noaa.gov/products/wesley/wgrib2/index.html

- **"Gribfilter"** streams data from NOMADS servers to clients. A web page interface filter is in place to allow the user, to select files by forecast time, for partial variable, access by region, or vertical level by variable. Users receive a re-packaged GRIB file streamed to the user workstation or the "URL commands" for program scripting. For details see:
  http://www.cpc.ncep.noaa.gov/products/wesley/scripting_grib_filter.html

NOMADS Ensemble Probability Tool

- **The Ensemble Probability Tool:** A web based client application that can be used to provide threshold information to the user. The program obtains the Global Ensemble Forecast information matrix from the server as described above from 400 separate binary packed GRIB2 files, and returns the information to the user as a probability display. A user then determines a proper threshold of their defined weather event. See demo at http://nomads.ncdc.noaa.gov/EnsProb/

- Users can determine their threshold (to tolerate false alarms) for a weather event alert, that is sent to email and cell phone (SMS) text-message when the threshold is met in a similar application http://nomads-cloud.net/cgi-bin/ensprob_wb4.pl.

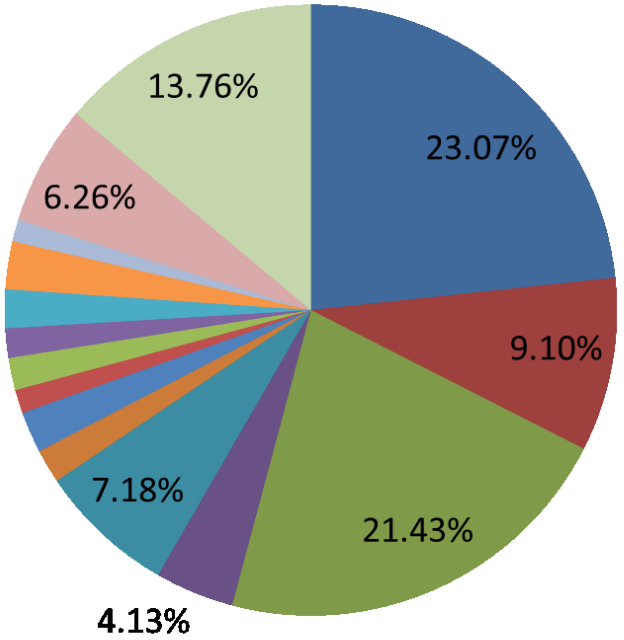
- The returned URL for this application will repeat the alert process, that is, sending user notification only if the threshold criteria is met, otherwise no action is taken, so the user may repeat the command in cron or a scheduler and assured of the latest forecast cycle (4X/day).

Access Statistics by Domain FY14

FY14 NCEI Users Access (Historical)

- ~1.2PB downloaded
- Distinct IP's: 110,000
- Successful Requests: ~165 million
  .com requests up 12%

Archive Holdings: 998.28 TB



NCEI: % of total volume

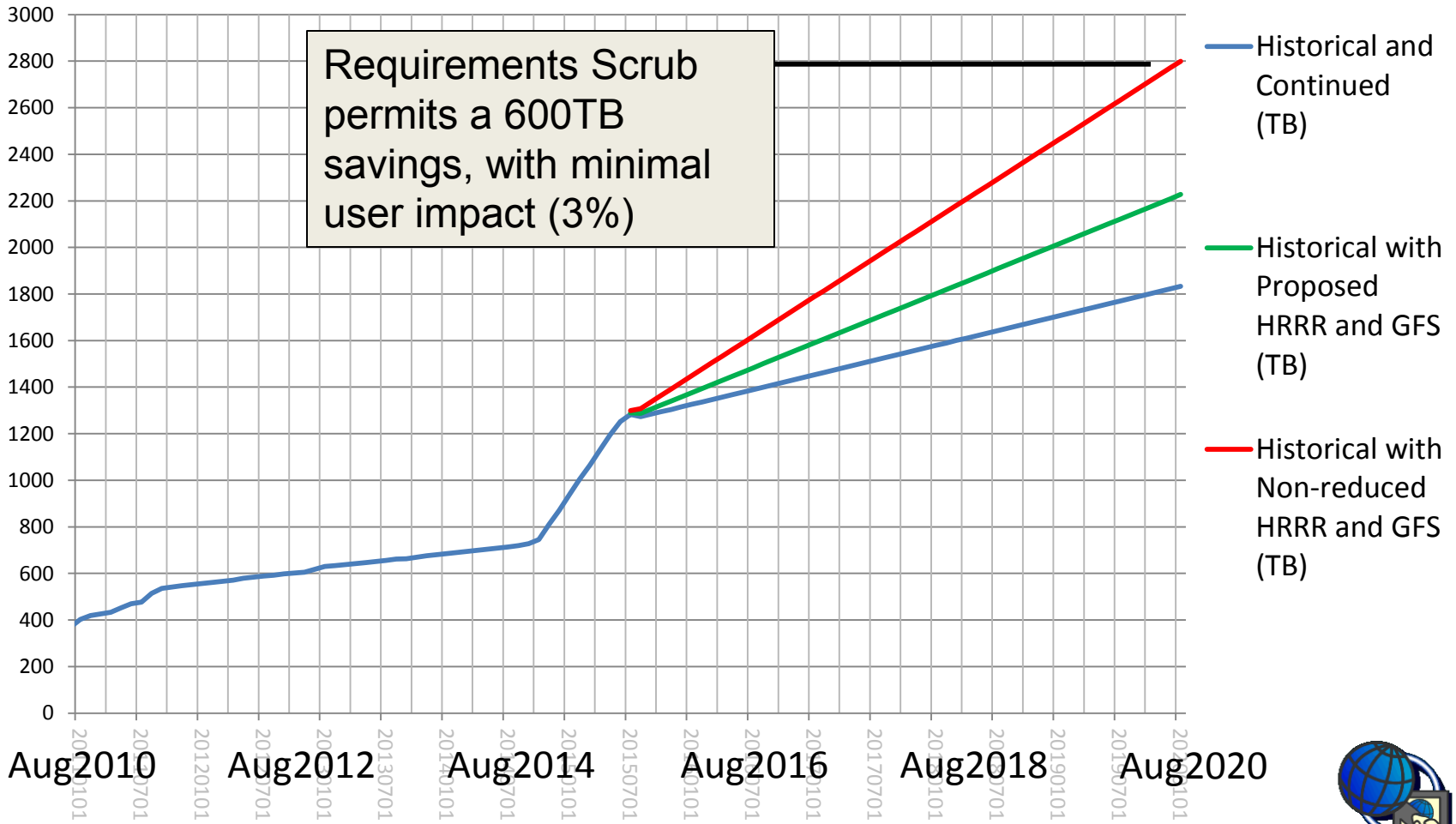
Real-Time NOMADS at NCEP:
~1.5PB downloaded in FY14

GO-ESSP

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

NCDC NOMADS Model Data Growth REDUCTION

Historical (Aug 2010 - Dec 2014) and Proposed (Jan 2014 - Jan 2020)
2 NEW MODELS with Reduced File set



GO-ESSP

Data Volume Considerations

Requested and Dynamic Dataset Scrubbing	
Dataset	**Growth (TB/year)**
HDSS Legacy	56
CFSv2 Op. R-T	70
HRRR All	153
HRRR wrfprs-only	**53**
GFS 0.25 Degree, all PGRB	58
GFS 0.25 Degree, pgrb2-only	**33**
NARR-Corrected (5TB)	**~**

New Potential Data (Static)		
Dataset	**Amount (TB)**	**Available**
20CRv2	104	2015
20CRv3	250	2016
20CRv4	4676+	2020
NMME Phase-I	0.5	2013
NMME Phase-II	>tbd	2015
Planned CFSv3	12,000	2018

Data Reduction/Compression *Analysis*
The Challenges of Exascale Computing

Data Type	Volume Growth Rate	Potential Data Reduction	"Compression"	Total Volume (PB)	
				FY18	FY30 est.
NWP and Ensembles[1]	Doubling/2 yrs*	Yes 5 Yrs, Del Fcsts	GRIB/NetCDF	6.4	~49,000
AOGCM's Climate[2]	Step Function	Yes Maintain 2 IPCC Cycles	NetCDF (4)	Tbd	~40,000
Reanalysis[3]	Step Function	Yes Maintain 2 Cycles Gaussian Grid Req?**	GRIB/NetCDF (4)	12.0+	~36,000
Regional Downscaled[4]	Linear	Yes 10 years	NetCDF (4)	Tbd	5,000
Reanalysis Clearinghouse Observations[5]	Linear	No WCRP/WOAP/other	Various	Reanalysis.org	1.00
Notes:				18.4+ PB	131 EB

1. Data Reduction: NWP forecasts removed after 5 years. Only Analysis and Initial Conditions retained.
2. Climate Models increasing as stewardship and user access requirements increase under NOAA.
3. NOAA reanalysis once performed every 5-10 years. New IT resources and NOAA "Clearinghouse" will allow on-going reanalysis.
4. Only critical regional and local downscaled analysis retained. Regional offices can maintain as req'ed.
5. NOAA recommendation to create a database of in-situ obs for on-going reanalysis of the climate system. Multi-Agency involvement: NOAA, NASA, ESCMF, EPA, NCAR, USGS, et al.; "Reanalysis.org".

* NWP Estimates based on verified linear to log exponential growth using empirical data since 1966 (J. Alpert 2004).

** Gaussian Grids consume high volumes with limited (but important users). Could users obtain directly from producer?

GO-ESSP

NOMADS FY15 Project Plans

Advances slow, but steady (significant budget reduction for NOMADS over the last two years); expected to continue. Prioritized to do list:

- Implementing a NCEI bulk- order and sub-setting services in the NOAA NESDIS archive activity: Comprehensive Large Array Stewardship System (CLASS) (800TB)
  - NCMP is the 1st "Archive as a Service" (AaaS) component in CLASS **(done)**

- Implementing a Center-wide GridFTP and Globus capability. **(Nearly done)**

  (Thanks Rachana and DOE's Argonne Nat'l Lab!).

- Earth System Grid Federation (ESGF) (thanks Dean and PCMDI and entire ESGF team (Luca et al.!!) Test node Q3
  - Plans include subsets of CFS, National Multi-Model Ensemble (NMME) GHCN, Reynolds sst's, CDR's. Supporting Obs4MIPS with NCEI flags datasets for NCEP Seasonal-to-Interannual fcst products.

FY15-FY16 Plans

- Install ERDDAP
  - New agile development teams at NCEI- leveraging Ocean-NOMADS resources to implement ERDDAP at NCEI
- Turn on the embedded LAS within ESGF
- Google Cloud and Earth Engine via Climate Data Initiative: Global RTOFS Ocean Model, and 0.25 degree GFS; and time series aggregations of Climate Fcst System (CFS)
- NOMADS Team is steward of the NCEI-wide TDS for access to Center-wide data holdings. (Thank you John Caron: grib aggregations

NOMADS Data Holdings

Data Availability:

- CFS dataset variables
- NOMADS Holdings (as of Dec2014)

GO-ESSP Vision / What Next?

- **British Atmospheric Data Centre**
  - Bryan Lawrence – Director, British Atmospheric Data Centre
- **Geophysical Fluid Dynamics Laboratory**
  - V. Balaji, Head, Modeling Group, Princeton/GFDL
- **The German Climate Computing Centre**
  - Michael Lautenschlager (NeRC Grid)
- **Lawrence Livermore National Laboratory**
  - Dean Williams, PCMDI, Chief Archive Services/CMIP5 , ESGF
- **National Center for Atmospheric Research**
  - Don Middleton, Senior Manager, Enabling Technologies, ESGF
- **Pacific Marine Environmental Laboratory**
  - Steve Hankin (retired) Unified Access Framework, DMIT
- **NOAA/Earth Systems Research Laboratory**
  - Cecelia Deluca (National Climate Projection and Prediction NCPP prototype)
- **NOAA/National Climatic Data Center**
  - Glenn Rutledge, (Program Manager NOMADS/NCMP)

http://go-essp.gfdl.noaa.gov/

Why GO-ESSP?

Why:

- Decision makers are seeking information to plan and respond to climate variability and change (e.g., agriculture, water, energy, sectors etc).

- Improved and coordinated access to model and other data via the GO-ESSP collaboration provides (me) with directed applications and resources to leverage and improve such access, and thus, for users a better understanding of climate phenomena including drought, El Niño, hydrologic cycle, etc.

- Coordinated GO-ESSP portal and ESGF access and archives permit community analysis & feedback to improve NWP, Seasonal-to-Inter-annual and Decadal simulations.

- Real-world developers, solving real-world problems
  - Closer link to users. This group has been doing this awhile…

The Value of GO-ESSP

- WWRP/WCRP/WIP and the CMIP process is providing outstanding services to the climate modeling community through ESGF services. They include advancements identified within GO-ESSP such as embedded apps and services like TDS, LAS, CMOR, etc., within ESGF.

- Limited resources dictate the need for (my team) leveraging via GO-ESSP !

- NOMADS NCEI glad to participate in GO-ESSP, and potentially host an upcoming meeting in Asheville.

- Larger conferences such as AGU/EGU/AMS etc. are just that …larger
  - Small directed settings such as GO-ESSP continues to be an extremely valuable resource to NOMADS and it's users (100's thousands in FY14!)
  - 12K+ attendees at AGU scare me…..

What NOMADS still needs to do…

- Improved support for assessments, downscaling, diagnostics, calibration of NWP ensembles and model-to-obs inter-comparisons, and aggregations of commonly used variables of reanalysis datasets.

- Regridding: Preserving the monthly mean, from high res to low res and back and forth.

- Format Conversions, Transformations & Geo-rectification: Grib1,Grib2 to CF complaint NetCDF; and from NetCDF3 to NetCDF4

- More general simpler tools for dumping data (of all formats) to CF compliant NetCDF (and we deal with BUFR and Grib!)

You get the point. Many of the talks at this GO-ESSP 2015 Workshop assist my Team to achieve some of these goals without reinventing.

Possible Next Steps for GO-ESSP?

- Continue this outstanding “small group” activity !!
  - This is not like AGU to me. Directed efforts & sharing is what works!
- PI’s conduct a strategy meeting
  - Roles/Resources/Goals. Conduct co-I (re)-scoping meeting with a call for community input
- Role for GO-ESSP
  - EOS (or maybe BAMS) article on GO-ESSP ?
  - Advance Computational Data Analysis
    - Don’t download: analyze on host
- Future of GO-ESSP
  - Keep GO-ESSP. Value in small group not big Conferences.
  - New (co-I) Partners? NASA, USGS, BOM, Canada, etc.

Thanks!

- QUESTIONS?
- Glenn.Rutledge@noaa.gov
- http://www.ncdc.noaa.gov/data-access/model-data

Member GO-ESSP

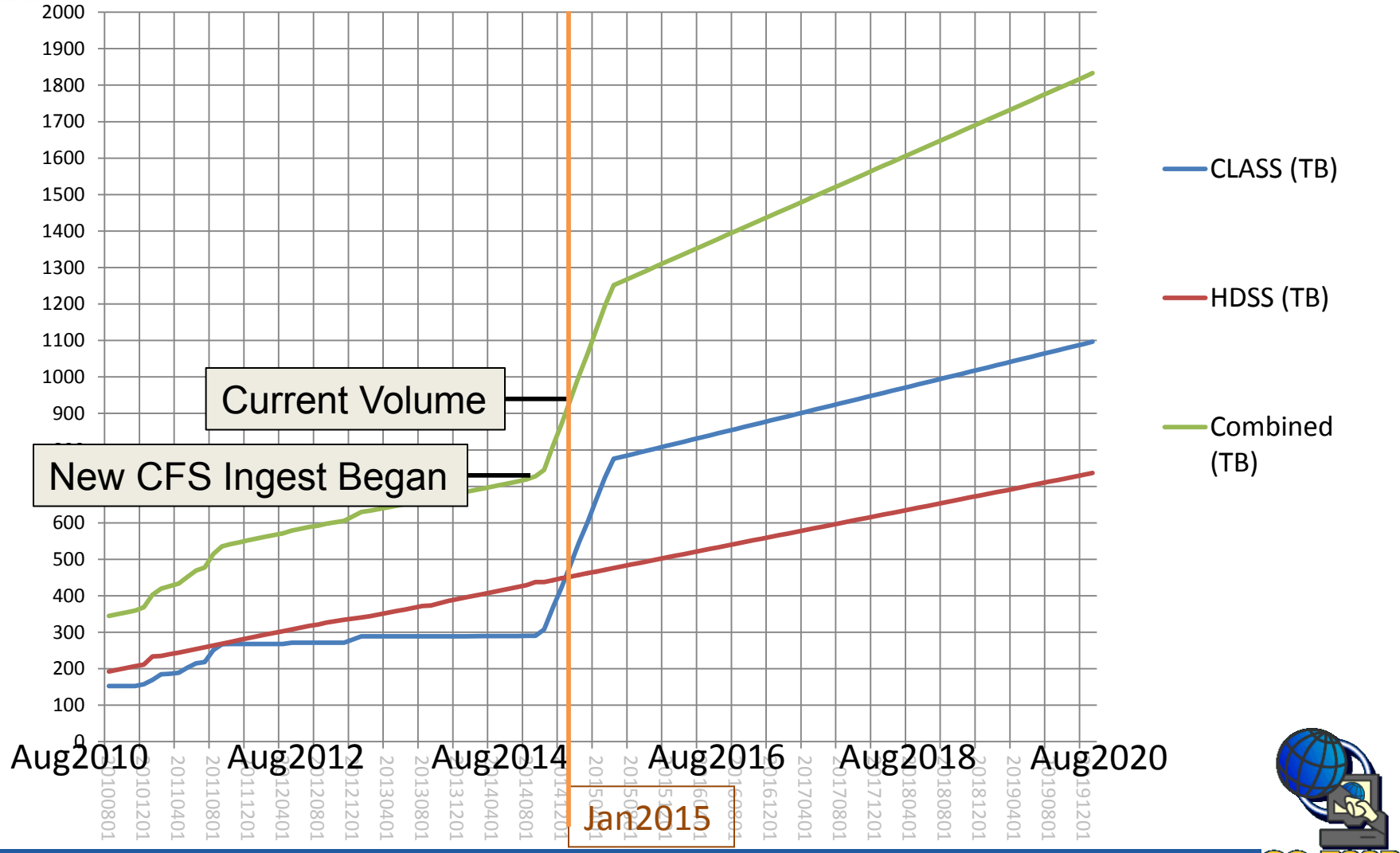
NOMADS
The NOAA Operational Model Archive and Distribution System

NOAA NCDC

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

NCDC NOMADS Model Data Growth

Historical (Aug 2010 - Dec 2014) and Continued (Jan 2014 - Jan 2020)
NO NEW DATA



GO-ESSP

What NOMADS still needs to do…

- Ratio of variances of (surface temperature)/(skin temperature)
- Radiative flux measures: TOA and surface SW and LW
- Examination of vertical structure of simulated Cloud fraction
- Investigate quantification of surface albedo and change
- Quantification of extreme events from models compared to observations (frequency and magnitude)
- Derived ENSO, NAO, indices compared to observations, emphasize timing of reoccurrence and magnitude.
- Simulation of satellite measures from model data

  ratio of variance of skin temp and simulated measure

Credit: Jay Hnilo DOE

What We Still Need…(can NOMADS leverage from GO-ESSP?)

ADVANCED / INTERMEDIATE USERS

• Incorporate first look diagnostic capabilities. These could be both a thumbnail plot, brief description, resultant data in NetCDF form and (initially) python source code and OGC compliant (ESRI) plots.

• Average Annual cycles, diurnal cycles, annual averages, examination of anomalies, measures of extreme events, climate sensitivity, decadal trends, ratio of variances, Empirical Orthogonal Functions and where applicable comparisons to observations.

• Examination of tropical variability, calculation of ENSO indices, simulation of satellite measures from both models and observations

What NOMADS still needs to do…

PROFESSIONAL USERS- Diagnostics & Model-to-Obs Inter-comparisons

- On-line access to Climate Model Analytical Engines
  - Climate model Data Analysis Tool (CDAT) for advanced model diagnostics and model-to-obs intercomparisons.
  - Develop easy to access thumbnail plots across GO-ESSP members generated from CDAT python code as to be extensible with CDAT to accurately geo-locate non-discrete points to grids.

- Variability (identification of regions of climate sensitivity)
  - Tools will be developed to find a simple ratio of variances for (projected model results) / (observations)
  - e.g., Twentieth Century Reanalysis Project, or Climate of the 20th Century run or, 2x or 4xCO2 GFDL IPCC CM.x runs

- Decadal Temperature Trends and Average Annual Cycles

Anomalies can be pre-staged. Average annual cycle differences between a control run and a doubled CO2 would infer that presently precip maximums are in June – but in a 2xCO2 run precip maximizes in January. These are important direct measures appropriate for an advanced audience.

Downscaling

DOWNSCALING
Close Coordination with U.S. National Assessment Program at USGCRP

Dynamical:
Initially use NCAR WRF for downscaling of global reanalysis, climate of the 20th century runs and 2xCO2 model runs. Output initially will emphasize state variables and those appropriate for study of the hydrological cycle and alternative energy production (e.g., wind, solar isolation).

Statistical:
NCMP will advance existing NCDC Sector Teams and work with mitigation, economic, and private sector researchers to formulate appropriate methods for application of this approach. We will emphasize tools to perform the calculations on the server side or give access to monthly-mean model output for data generation by the researchers themselves. It is thought ASCII text output would be sufficient for this method in the short term. (Heyhoe and Boyle guidance)

NOMADS - NCMP
Eg. Science Tools for Users

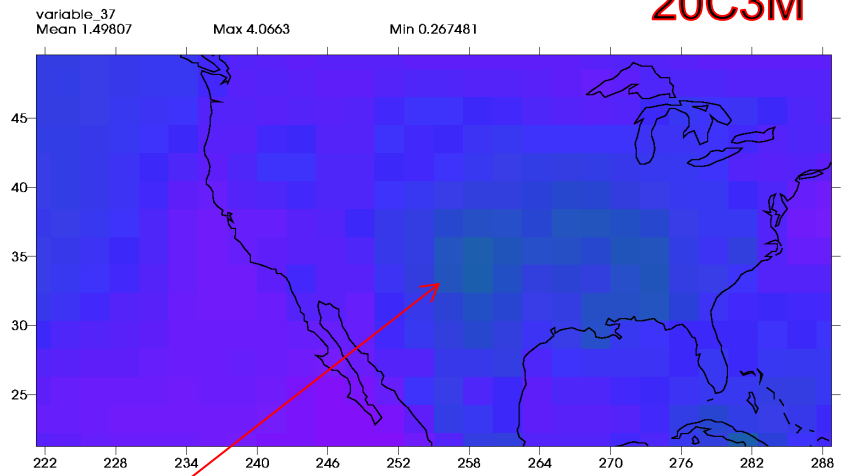
Climate of the 20th century runs (Annual averages):
Global variance of surface temperature (C).

NCEP/NCAR R2 reanalysis (1948-2000) = 0.064
GFDL 2.1 (1948-2000) = 0.071

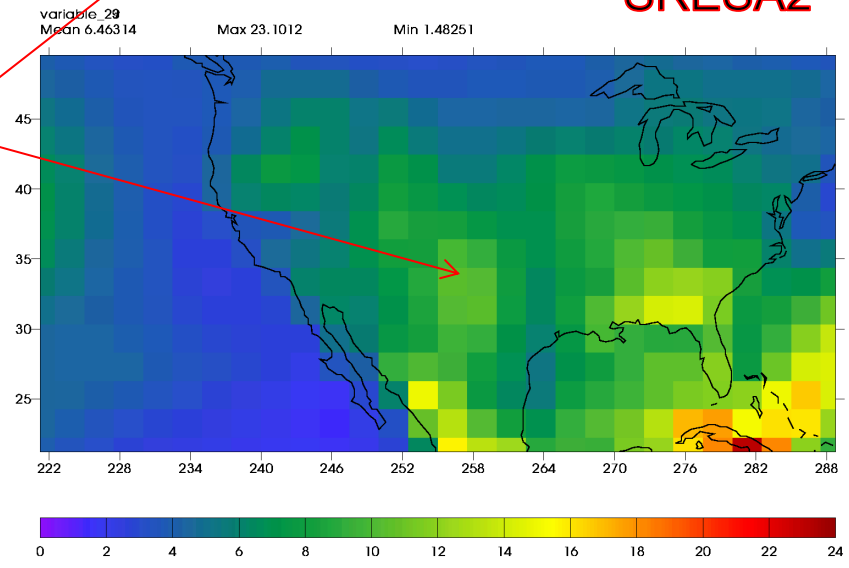
GFDL 2.1 (2000-2099) A2 = 0.841
GFDL 2.1 (2000-2099) A1B = 0.553
GFDL 2.1 (2000-2099) B1 = 0.193

A ratio of variance of annually averaged surface temperature (same scale)

20C3M



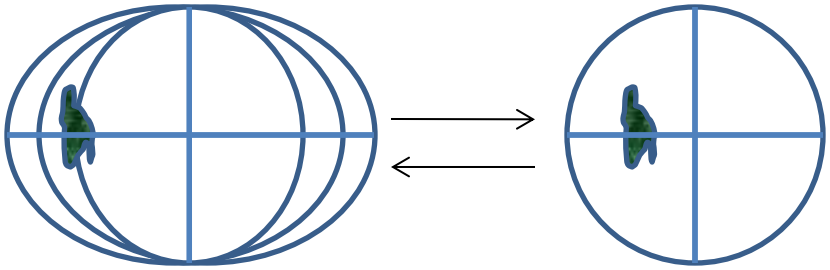
SRESA2



Credit: Jay Hnilo

NOMADS - NCMP
Support for GIS Users

Oblate spheroid (aka, Earth) to sphere

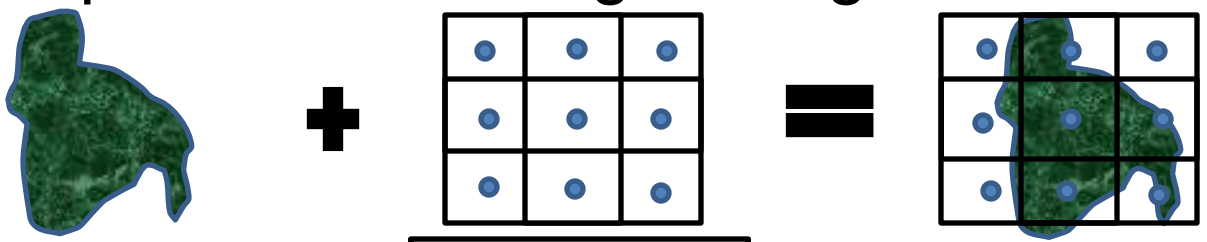


EARTH MODEL

At present there are no commonly used transformations when going from one to the other

No Datums included in datasets

Shapefile defined region to gridded data

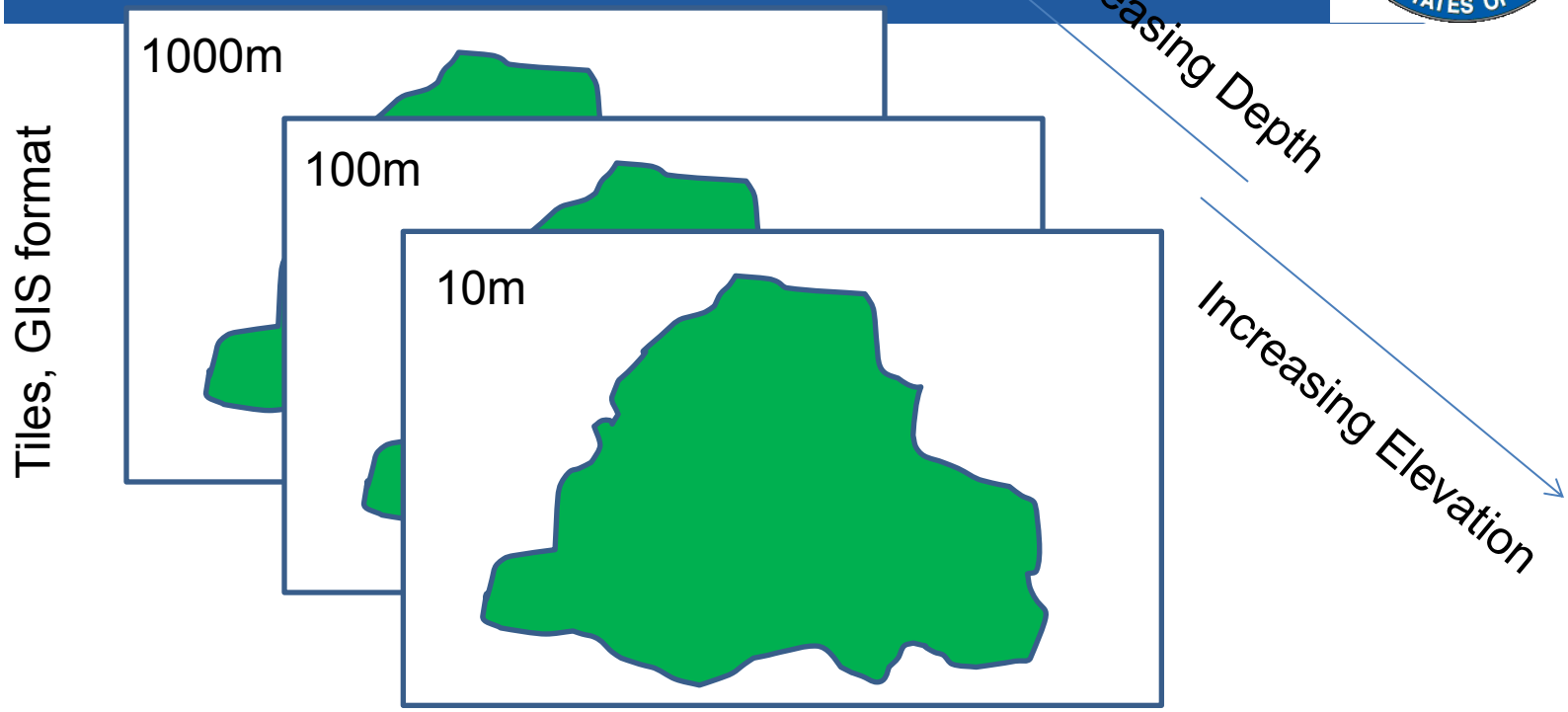


Shapefiles: Bound Regions

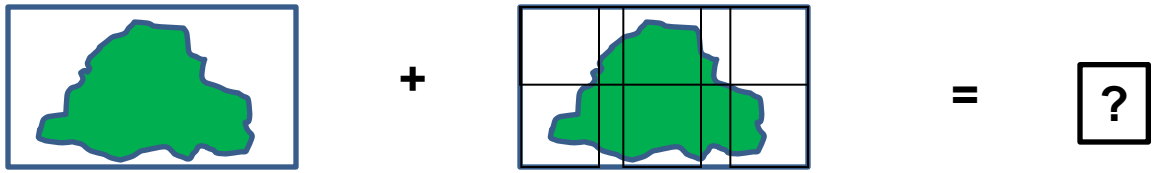
Grids defined by Lat & Lon Bounds

?

Advanced Tasks: GIS to NetCDF



Gridding high/deep topography to netCDF



GIS geometries

Tiles at different scales